OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844

DATE: 07/20/2001
TIME: 11:50:47

Input Set : N:\Crf3\RULE60\09897844.txt
Output Set: N:\CRF3\07202001\I897844.raw

ENTERED

```
  3 <110> APPLICANT: Cox III, George Norbert
  4       Case, Casey Christopher
  5       Eisenberg, Stephen P.
  6       Jarvis, Eric Edward
  7       Spratt, Sharon Kaye
  8       Sangamo Biosciences, Inc.
 10 <120> TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells Using
 11       Zinc Finger Proteins
 13 <130> FILE REFERENCE: 019496-002200US
 15 <140> CURRENT APPLICATION NUMBER: 09/897,844
 16 <141> CURRENT FILING DATE: 2001-07-02
 18 <150> PRIOR APPLICATION NUMBER: 09/229,037
 19 <151> PRIOR FILING DATE: 1999-01-12
 21 <160> NUMBER OF SEQ ID NOS: 40
 23 <170> SOFTWARE: PatentIn Ver. 2.0
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 25
 27 <212> TYPE: PRT
 28 <213> ORGANISM: Artificial Sequence
 30 <220> FEATURE:
 31 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif
 32       of C2H2 class of zinc finger proteins (ZFP)
 34 <220> FEATURE:
 35 <221> NAME/KEY: MOD_RES
 36 <222> LOCATION: (2)..(3)
 37 <223> OTHER INFORMATION: Xaa = any amino acid
 39 <220> FEATURE:
 40 <221> NAME/KEY: MOD_RES
 41 <222> LOCATION: (4)..(5)
 42 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
 44 <220> FEATURE:
 45 <221> NAME/KEY: MOD_RES
 46 <222> LOCATION: (7)..(18)
 47 <223> OTHER INFORMATION: Xaa = any amino acid
 49 <220> FEATURE:
 50 <221> NAME/KEY: MOD_RES
 51 <222> LOCATION: (20)..(22)
 52 <223> OTHER INFORMATION: Xaa = any amino acid
 54 <220> FEATURE:
 55 <221> NAME/KEY: MOD_RES
 56 <222> LOCATION: (23)..(24)
 57 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
 59 <400> SEQUENCE: 1
W--> 60 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
     61   1               5                  10                  15
W--> 63 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His
     64              20                  25
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844

DATE: 07/20/2001
TIME: 11:50:47

Input Set : N:\Crf3\RULE60\09897844.txt
Output Set: N:\CRF3\07202001\I897844.raw

```
    67 <210> SEQ ID NO: 2
    68 <211> LENGTH: 10
    69 <212> TYPE: DNA
    70 <213> ORGANISM: Artificial Sequence
    72 <220> FEATURE:
    73 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
    74       with two overlapping D-able subsites
    76 <220> FEATURE:
    77 <221> NAME/KEY: modified_base
    78 <222> LOCATION: (1)..(2)
    79 <223> OTHER INFORMATION: n = g,a,c or t
    81 <220> FEATURE:
    82 <221> NAME/KEY: modified_base
    83 <222> LOCATION: (5)
    84 <223> OTHER INFORMATION: n = g,a,c or t
    86 <220> FEATURE:
    87 <221> NAME/KEY: modified_base
    88 <222> LOCATION: (8)
    89 <223> OTHER INFORMATION: n = g,a,c or t
    91 <220> FEATURE:
    92 <221> NAME/KEY: modified_base
    93 <222> LOCATION: (9)
    94 <223> OTHER INFORMATION: n = a,c or t; if g, then position 10 cannot be g
    95       or t
    97 <220> FEATURE:
    98 <221> NAME/KEY: modified_base
    99 <222> LOCATION: (10)
   100 <223> OTHER INFORMATION: n = a or c; if g or t, then position 9 cannot be g
   102 <400> SEQUENCE: 2                                                      10
W--> 103 nngkngknnn
   106 <210> SEQ ID NO: 3
   107 <211> LENGTH: 10
   108 <212> TYPE: DNA
   109 <213> ORGANISM: Artificial Sequence
   111 <220> FEATURE:
   112 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
   113       with three overlapping D-able subsites
   115 <220> FEATURE:
   116 <221> NAME/KEY: modified_base
   117 <222> LOCATION: (1)..(2)
   118 <223> OTHER INFORMATION: n = g,a,c or t
   120 <220> FEATURE:
   121 <221> NAME/KEY: modified_base
   122 <222> LOCATION: (5)
   123 <223> OTHER INFORMATION: n = g,a,c or t
   125 <220> FEATURE:
   126 <221> NAME/KEY: modified_base
   127 <222> LOCATION: (8)
   128 <223> OTHER INFORMATION: n = g,a,c or t
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844

DATE: 07/20/2001
TIME: 11:50:47

Input Set : N:\Crf3\RULE60\09897844.txt
Output Set: N:\CRF3\07202001\I897844.raw

```
    130 <400> SEQUENCE: 3                                                  10
W--> 131 nngkngkngk
    134 <210> SEQ ID NO: 4
    135 <211> LENGTH: 5
    136 <212> TYPE: PRT
    137 <213> ORGANISM: Artificial Sequence
    139 <220> FEATURE:
    140 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
    142 <400> SEQUENCE: 4
    143 Asp Gly Gly Gly Ser
    144   1               5
    147 <210> SEQ ID NO: 5
    148 <211> LENGTH: 5
    149 <212> TYPE: PRT
    150 <213> ORGANISM: Artificial Sequence
    152 <220> FEATURE:
    153 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
    155 <400> SEQUENCE: 5
    156 Thr Gly Glu Lys Pro
    157   1               5
    160 <210> SEQ ID NO: 6
    161 <211> LENGTH: 9
    162 <212> TYPE: PRT
    163 <213> ORGANISM: Artificial Sequence
    165 <220> FEATURE:
    166 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
    168 <400> SEQUENCE: 6
    169 Leu Arg Gln Lys Asp Gly Glu Arg Pro
    170   1               5
    173 <210> SEQ ID NO: 7
    174 <211> LENGTH: 4
    175 <212> TYPE: PRT
    176 <213> ORGANISM: Artificial Sequence
    178 <220> FEATURE:
    179 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
    181 <400> SEQUENCE: 7
    182 Gly Gly Arg Arg
    183   1
    186 <210> SEQ ID NO: 8
    187 <211> LENGTH: 5
    188 <212> TYPE: PRT
    189 <213> ORGANISM: Artificial Sequence
    191 <220> FEATURE:
    192 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
    194 <400> SEQUENCE: 8
    195 Gly Gly Gly Gly Ser
    196   1               5
    199 <210> SEQ ID NO: 9
    200 <211> LENGTH: 8
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844

DATE: 07/20/2001
TIME: 11:50:47

Input Set : N:\Crf3\RULE60\09897844.txt
Output Set: N:\CRF3\07202001\I897844.raw

```
201 <212> TYPE: PRT
202 <213> ORGANISM: Artificial Sequence
204 <220> FEATURE:
205 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
207 <400> SEQUENCE: 9
208 Gly Gly Arg Arg Gly Gly Gly Ser
209   1               5
212 <210> SEQ ID NO: 10
213 <211> LENGTH: 9
214 <212> TYPE: PRT
215 <213> ORGANISM: Artificial Sequence
217 <220> FEATURE:
218 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
220 <400> SEQUENCE: 10
221 Leu Arg Gln Arg Asp Gly Glu Arg Pro
222   1               5
225 <210> SEQ ID NO: 11
226 <211> LENGTH: 12
227 <212> TYPE: PRT
228 <213> ORGANISM: Artificial Sequence
230 <220> FEATURE:
231 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
233 <400> SEQUENCE: 11
234 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
235   1               5                  10
238 <210> SEQ ID NO: 12
239 <211> LENGTH: 16
240 <212> TYPE: PRT
241 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
246 <400> SEQUENCE: 12
247 Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
248   1               5                  10                  15
251 <210> SEQ ID NO: 13
252 <211> LENGTH: 25
253 <212> TYPE: DNA
254 <213> ORGANISM: Artificial Sequence
256 <220> FEATURE:
257 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
258       region surrounding initiation site of vascular
259       endothelial growth factor (VEGF) gene containing
260       two 9-base pair target sites
262 <220> FEATURE:
263 <221> NAME/KEY: protein_bind
264 <222> LOCATION: (4)..(12)
265 <223> OTHER INFORMATION: upstream 9-base pair ZFP VEGF1 target site
267 <220> FEATURE:
268 <221> NAME/KEY: protein_bind
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844

DATE: 07/20/2001
TIME: 11:50:47

Input Set : N:\Crf3\RULE60\09897844.txt
Output Set: N:\CRF3\07202001\I897844.raw

```
269 <222> LOCATION: (14)..(22)
270 <223> OTHER INFORMATION: downstream 9-base pair ZFP VEGF3a target site
272 <400> SEQUENCE: 13
273 agcggggagg atcgcggagg cttgg                                          25
276 <210> SEQ ID NO: 14
277 <211> LENGTH: 298
278 <212> TYPE: DNA
279 <213> ORGANISM: Artificial Sequence
281 <220> FEATURE:
282 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP
283       construct targeting upstream 9-base pair target
284       site in VEGF promoter
286 <220> FEATURE:
287 <221> NAME/KEY: CDS
288 <222> LOCATION: (2)..(298)
289 <223> OTHER INFORMATION: VEGF1
291 <400> SEQUENCE: 14
292 g gta ccc ata cct ggc aag aag aag cag cac atc tgc cac atc cag ggc 49
293   Val Pro Ile Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly
294     1               5                  10                  15
296 tgt ggt aaa gtt tac ggc aca acc tca aat ctg cgt cgt cac ctg cgc   97
297 Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg
298             20                  25                  30
300 tgg cac acc ggc gag agg cct ttc atg tgt acc tgg tcc tac tgt ggt  145
301 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly
302         35                  40                  45
304 aaa cgc ttc acc cgt tcg tca aac ctg cag cgt cac aag cgt acc cac  193
305 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His
306     50                  55                  60
308 acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg  241
309 Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met
310 65                  70                  75                  80
312 cgt agt gac cac ctg tcc cgt cac atc aag acc cac cag aat aag aag  289
313 Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys
314                 85                  90                  95
                                                                    298
316 ggt gga tcc
317 Gly Gly Ser
320 <210> SEQ ID NO: 15
321 <211> LENGTH: 99
322 <212> TYPE: PRT
323 <213> ORGANISM: Artificial Sequence
325 <220> FEATURE:
326 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP
327       construct targeting upstream 9-base pair target
328       site in VEGF promoter
330 <400> SEQUENCE: 15
331 Val Pro Ile Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly
332   1               5                  10                  15
334 Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/897,844

DATE: 07/20/2001
TIME: 11:50:48

Input Set : N:\Crf3\RULE60\09897844.txt
Output Set: N:\CRF3\07202001\I897844.raw

L:60 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:63 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:103 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:131 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3